Size Distribution of Amyloid Nanofibrils

Raffaela Cabriolu,[†] Dimo Kashchiev,[‡] and Stefan Auer[†*]
[†]Centre for Molecular Nanoscience, School of Chemistry, University of Leeds, Leeds, United Kingdom; and [‡]Institute of Physical Chemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria

ABSTRACT We consider the size distribution of amyloid nanofibrils (protofilaments) in nucleating protein solutions when the nucleation process occurs by the mechanism of direct polymerization of β-strands (extended peptides or protein segments) into β-sheets. Employing the atomistic nucleation theory, we derive a general expression for the stationary size distribution of amyloid nanofibrils constituted of successively layered β-sheets. The application of this expression to amyloid $\beta_{1\text{-}40}$ (Aβ_{40}) fibrils allows us to determine the nanofibril size distribution as a function of the protein concentration and temperature. The distribution is most remarkable with its exhibiting a series of peaks positioned at "magic" nanofibril sizes (or lengths), which are due to deep local minima in the work for fibril formation. This finding of magic sizes or lengths is consistent with experimental results for the size distribution of aggregates in solutions of Aβ_{40} proteins. Also, our approach makes it possible to gain insight into the effect of point mutations on the nanofibril size distribution, an effect that may play a role in experimentally observed substantial differences in the fibrillation lag-time of wild-type and point-mutated amyloid-β proteins.

INTRODUCTION

The assembly of proteins into amyloid fibrils is an important phenomenon with wide implications ranging from human disease to nanoscience. Amyloid fibril formation is associated with a growing number of neurological and systemic diseases (1) including Alzheimer's and Parkinson's diseases, and the application of amyloid fibrils as molecular building blocks in biosensors, tissue engineering, and antibacterial agents has been demonstrated (2). The main insight into the structures formed by proteins comes from structural biology experiments. Such experiments are very challenging for protein aggregates because of the aggregate transient nature and structural heterogeneity (1). An essential point in our current understanding is that peptides and proteins unrelated in sequence and structure have been shown to convert into large fibrillar aggregates including amyloid fibrils (1). These aggregates share a common cross-β structure formed by intertwined layers of β-sheets extending in a direction parallel to the fibril elongation axis (3–8). Oligomeric aggregates are often found as precursors of amyloid fibrils (9–11) and their structure can vary from that of highly disordered aggregates that are composed of unstructured proteins to that of nativelike aggregates in which the proteins retain their native state.

It is now well established (9,10,12–30) that fibrillar protein aggregates can form by a nucleation mechanism. Time-resolved optical experiments that measure the fluorescence signal arising from a dye-binding molecule such as thioflavin-T bound to the protein aggregate have been used to determine the lag time, during which no aggregates are detected, and the maximal rate of conversion of protein monomers into aggregates. Experiments (28,31–41) have been performed to reveal the relationship between the physicochemical properties of the natural amino acids and the kinetics of amyloid formation. These experiments have in turn been used to substantiate phenomenological models (42–46) able to predict changes in the propensity of proteins to aggregate upon mutation as well as to identify amino-acid sequences of proteins that are likely to belong to the fibril core. To gain more insight into the nucleation mechanism and to determine important nucleation parameters such as the fibril nucleus size and nucleation rate, it is necessary to monitor the time evolution of the aggregates formed in a nucleating protein solution. Recently, small-angle x-ray scattering (47,48) and fluorescence correlation spectroscopy experiments (29,49) have been used to monitor the population, structural changes, and size of protein aggregates in the solution. Understanding and interpreting such experiments requires developing a theoretical framework describing amyloid fibril nucleation.

In previous studies, we used concepts from the theory of overall crystallization to describe the kinetics of overall protein aggregation (50), and applied the classical and atomistic nucleation theories (CNT and ANT, respectively) to treat the nucleation of amyloid fibrils when the process takes place by the mechanism of direct polymerization of single β-strands (extended peptides or protein segments) into β-sheets (51,52). In this, our treatment differed from that of Schmit et al. (53), who considered the case of amyloid fibril nucleation in two steps: single β-strands assemble into disordered oligomers, which then transform into β-sheets. The application of ANT has been demonstrated by correlating theoretical fibril nucleation rates with experimental lag times in the fibrillation kinetics of β_2-microglobulin and amyloid $\beta_{1\text{-}40}$ (Aβ_{40}) proteins (54). The objective of this study is to employ nucleation theory for

Submitted September 13, 2011, and accepted for publication September 30, 2011.

*Correspondence: s.auer@leeds.ac.uk

Editor: R. Astumian.

0006-3495/11/11/2232/10 $2.00

doi: 10.1016/j.bpj.2011.09.053

determination of the size distribution of amyloid nanofibrils (protofilaments) in protein solutions when fibril nucleation occurs in one step by polymerization of β-strands from the solution directly into β-sheets. We derive a general expression for the stationary size distribution of amyloid nanofibrils constituted of successively layered β-sheets and illustrate how this expression can be applied to the smallest $A\beta_{40}$ protofilaments. The results obtained are applicable to homogeneous nucleation that takes place when the protein solution is sufficiently pure and/or strongly supersaturated.

THEORY

Modeling the amyloid nanofibril

Nucleation of nanosized amyloid fibrils by direct polymerization of β-strands into β-sheets is operative when no fibril precursors such as dropletlike oligomers are formed as a first step in the fibril nucleation process. Fig. 1 defines the nanofibril and illustrates our model (52,54). The model is based on reported structural and morphological results for amyloid fibrils and microcrystals (3–8). Due to their strong hydrogen bonds, the virtually fully extended protein β-strands can arrange themselves into β-sheets that stretch forth along the fibril lengthening axis (the m axis in Fig. 1). The much weaker hydrophobicity-mediated interstrand bonds cause the β-strands to assemble along the fibril thickening axis (the i axis in Fig. 1) in such a way that a nanosized amyloid fibril composed of successively layered β-sheets can form. Because the fibril width is fixed and equal to the β-strand length, the fibril can be considered as a two-dimensional aggregate in the m,i plane, with building blocks (the β-strands) arranged in a two-dimensional lattice with simple rectangular symmetry (Fig. 1). Inasmuch as these blocks play the role of the atoms in the original ANT, with the term "atomistic" referring to the β-strands, this theory is readily applicable to our nanofibril model.

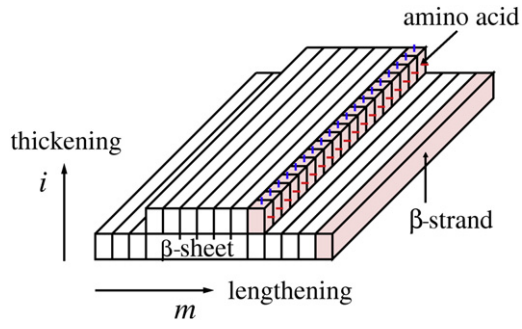

FIGURE 1 Schematic of nanosized fibril with thickness of two differently long β-sheets. The first and the second β-sheets are with length of 14 and 7 β-strands, respectively. Each of the 21 β-strands of the fibril is composed of 19 amino acids. The 19 red and 19 blue lines shown for the rightmost β-strand of the second β-sheet visualize the amino-acid strong and weak broken bonds that are in direction of the fibril lengthening and thickening (m and i) axes, respectively.

Sequence-specific binding energies

In the atomistic approach to nucleation of amyloid fibrils (52,54,55), important parameters are the binding energies of the β-strands within the fibril. We denote by E (J) and E_h (J) the binding energies between two nearest-neighbor β-strands along the m and i axes in Fig. 1, respectively. The excess energy of any n-sized nanofibril equals the fibril total surface energy (51) that, according to ANT (52,55), is given by the total energy $l_nE/2 + l_{h,n}E_h/2$ of the nearest-neighbor broken-bonds at the periphery of the fibril cross section in the m,i plane. In the above sum, n is the number of β-strands in the fibril, and l_n or $l_{h,n}$ is the number of all nearest-neighbor broken bonds that are parallel to the fibril m axis or i axis, respectively.

To calculate E and E_h, we assume that each amino acid in a β-strand (19 such amino acids are illustrated in Fig. 1 by the prisms in the rightmost β-strand of the shorter β-sheet) forms hydrogen bonds solely along the m axis, and that hydrophobicity-mediated bonds can be formed along both the m and i axes. Assuming further that bonds can only be formed between nearest-neighbor amino acids, the interstrand binding energies E and E_h are obtained by summation of the bond energies of such amino acids over all amino-acid pairs in neighboring β-strands:

$$E = \sum_{p} \left[\varepsilon_{jq}(p) + \varepsilon_{h,jq}(p)\right], \tag{1}$$

$$E_h = \sum_{p_h} \varepsilon_{h,jq}(p_h). \tag{2}$$

In Eq. 1, the summation goes over all amino-acid pairs p in two nearest-neighbor β-strands of a β-sheet, and in Eq. 2 it goes over all amino-acid pairs p_h in two nearest-neighbor β-strands of two successive β-sheets. The energies of a nearest-neighbor hydrogen amino-acid bond and a nearest-neighbor, hydrophobicity-mediated amino-acid bond between amino acids j and q are denoted by ε_{jq} (J) and $\varepsilon_{h,jq}$ (J), respectively ($\varepsilon_{jq} >> \varepsilon_{h,jq}$, because the hydrogen bond is much stronger than the hydrophobicity-mediated one).

The assignment of amino-acid, sequence-specific values to the binding energies ε_{jq} and $\varepsilon_{h,jq}$ allows application of our nanofibril model to concrete protein systems. In this study, the sequence-specific hydrogen-bond energies used are based on a statistical analysis (45) of the frequency that two residue types are found paired in neighboring β-strands within β-sheets in globular proteins. The pairing energy depends on the orientation (parallel or antiparallel) of the β-strands. We denote by $\varepsilon^P{}_{jq}$ and $\varepsilon^A{}_{jq}$ the dimensionless pairing energies for the parallel and antiparallel orientation, respectively, and employ the $\varepsilon^P{}_{jq}$ and $\varepsilon^A{}_{jq}$ values from Table 1 of Trovato et al. (45) to modulate the basic unit ε (J) of the hydrogen bond energy in such a way that the values of the binding energy ε_{jq} lie in the interval $[0, 2\varepsilon]$:

$$\varepsilon_{jq} = \frac{2\left(\varepsilon^{A}_{PRO,ASP} - \varepsilon^{P(A)}_{jq}\right)}{\left(\varepsilon^{A}_{PRO,ASP} - \varepsilon^{A}_{CYS,CYS}\right)}\varepsilon. \quad (3)$$

Here $\varepsilon^{A}_{CYS,CYS} = -2.57$ and $\varepsilon^{A}_{PRO,ASP} = 3.55$ are, respectively, the dimensionless energies of a CYS-CYS pair and a PRO-ASP pair in an antiparallel β-sheet, and $\varepsilon^{P(A)}_{jq}$ equals ε^{P}_{jq} or ε^{A}_{jq} for the parallel or antiparallel orientation. For the basic unit of the hydrogen bond we use $\varepsilon = 6.95 \times 10^{-21}$ J (1 kcal/mol), a value in the range of the hydrogen-bond energies measured experimentally (56).

Nanofibril evolution mechanism

The determination of the nanofibril size distribution requires specification of the sequence of shapes that the fibril takes during its evolution from the monomer size $n = 1$ to the considered size $n > 1$. In reality, there are many such sequences, but we choose a sequence of shapes that have the fibril equilibrium shape as a reference low-energy shape, because the latter corresponds to the minimal fibril total surface energy (51).

In this shape sequence (Fig. 2), the fibril initially evolves as a single β-sheet and its transition from one-dimensional to two-dimensional aggregate occurs at transition size n_t that is assumed to be given by the CNT formula (51) $n_t = E/E_h$. This size is an important parameter, because it characterizes the fibril equilibrium shape, i.e., the fibril thermodynamically favored aspect (length/thickness) ratio. Thus, the $(n_t + 1)$-sized fibril (the one with $n = 15$ in Fig. 2) has the shape of 1β-sheet of length n_t with one protein monomer (β-strand) adsorbed on the sheet surface. This monomer gives birth of the fibril second β-sheet and creates two kink sites at the fibril surface (marked by *plus signs* in Fig. 2). As shown in Cabriolu et al. (52), in a certain supersaturation range the $(n_t + 1)$-sized fibril requires maximum work for its formation and is therefore the so-called nucleus. Further monomer attachment to a kink site does not increase the fibril total surface energy, because the fibrils of size $n = n_t + 1, n_t + 2, \ldots, 2n_t$ have the same total number of broken bonds (see the fibrils with $n = 15$, 16, and 28 in Fig. 2).

Because from $n = 2n_t$ to $n = 2 \times 2n_t$ the CNT equilibrium shape requires fibril thickness between one and two β-sheets, the attachment of the next protein monomer to the $2n_t$-sized fibril can be assumed to occur at one of the two fibril ends. This lengthwise monomer attachment creates a kink site at the end of the fibril of size $n = 2n_t + 1$ (the fibril with $n = 29$ in Fig. 2) and the next monomer is attached to this site with no surface-energy cost. Thus, further fibril growth occurs by alternating creation and annihilation of kink sites at the fibril ends until the fibril reaches the size $n = 2 \times 2n_t$ (see the fibril with $n = 56$ in Fig. 2). At this size, the attachment of the next protein monomer can be assumed to occur again to the surface of one of the fibril two β-sheets, because the CNT equilibrium shape requires fibril thickness between two and three β-sheets for fibril sizes from $n = 2 \times 2n_t + 1$ to $n = 3 \times 3n_t$. The fibril of size $2 \times 2n_t + 1$ (the one with $n = 57$ in Fig. 2) will therefore have the shape of a double β-sheet, i.e., a 2β-sheet, of length $2n_t$ with one protein monomer adsorbed on the 2β-sheet surface rather than attached to one of the 2β-sheet ends. The $(2 \times 2n_t + 1)$-sized fibril is also remarkable, because it is the nucleus in a corresponding supersaturation range (52). The first monomer on the

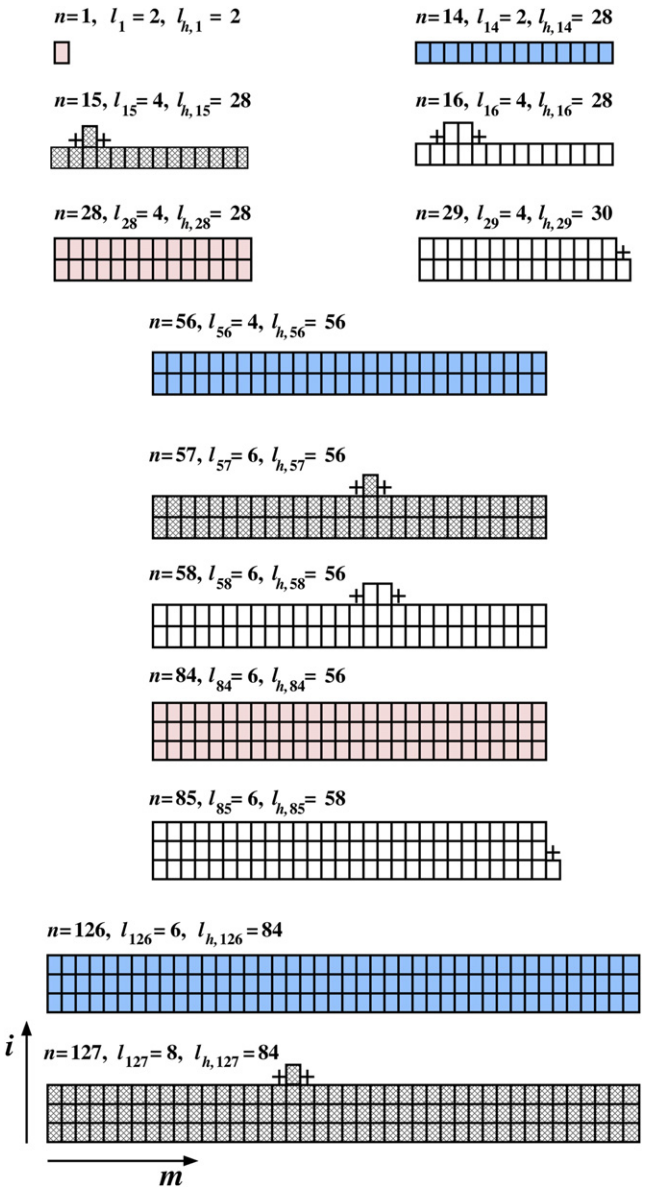

FIGURE 2 Sequence of nanofibril shapes corresponding to minimum total surface energy, which depicts the fibril evolution when the transition from one-dimensional to two-dimensional aggregate is at size $n_t = E/E_h = 14$. The first three nanofibrils of size n'_j (*pink*) and n''_j (*blue*) are shown. (*Hatched*) First three fibril nuclei. (*Pluses*) Kink sites at which β-strand attachment occurs at no surface-energy cost.

2β-sheet triggers the growth of the fibril third β-sheet, because it generates two kink sites at which subsequent monomers can attach themselves without changing the total number of the fibril broken bonds. In Fig. 2, the fibrils with $n = 57$ and 58 illustrate this situation.

When the third β-sheet is completed at size $3 \times 2n_t$, the fibril growth continues by lengthwise attachment of the next protein monomer (see the fibrils with $n = 84$ and 85 in Fig. 2). Subsequent lengthwise attachment of monomers to the fibril is thermodynamically favored only up to size $n = 3 \times 3n_t$ (that of the fibril with $n = 126$ in Fig. 2), because between this size and the size $n = 4 \times 4n_t$, to keep its shape close to the CNT equilibrium shape, the fibril has to increase its thickness by one more β-sheet. Thus, it can again be assumed that the fibril of size $n = 3 \times 3n_t + 1$ (which is the nucleus size in a respective supersaturation range (52)) will have the shape of a triple β-sheet, i.e., a 3β-sheet, of length $3n_t$ with one monomer adsorbed on the surface of the fibril outer β-sheets (see the fibril with $n = 127$ in Fig. 2). Further growth of the fibril can be expected to occur in the above manner provided it is not disturbed by other processes, e.g., fibril flocculation.

Nanofibril size distribution

As known from nucleation theory (e.g., Kashchiev (57)), the stationary size distribution of the aggregates of a nucleating single-component phase can be presented exactly as ($n = 1,2,3,\ldots$)

$$X_n = C_n \frac{\sum_{k=n}^{M-1}\left(\frac{1}{f_k C_k}\right)}{\sum_{k=1}^{M-1}\left(\frac{1}{f_k C_k}\right)}, \quad (4)$$

when nucleation takes place by the Szilard-Farkas mechanism, according to which the aggregates change size by random attachment and detachment solely of monomeric building units. Here n (or k) is the number of building units in an aggregate, X_n (m^{-3}) is the stationary concentration of n-sized aggregates, C_n (m^{-3}) is the respective equilibrium concentration of such aggregates, M is the number of building units in a large enough supernucleus aggregate, and f_n (s^{-1}) is the frequency of building-unit attachment to an n-sized aggregate.

Equation 4 is readily applicable to nucleation of amyloid fibrils in the case when the process occurs by the mechanism of direct polymerization of β-strands into β-sheets. Then n is the number of β-strands in an n-sized amyloid nanofibril and, according to ANT, C_n and f_n are of the form (52)

$$C_n = C_1 \exp[(n-1)s - (l_n - 2)\psi - (l_{h,n} - 2)\psi_h], \quad (5)$$

$$f_n = \frac{1}{2} f_1 l_n, \quad (6)$$

when only the nearest-neighbor interactions between the β-strands are taken into account.

In Eqs. 5 and 6, C_1 (m^{-3}) is the actual concentration of single β-strands in the protein solution, and f_1 (s^{-1}) is the frequency of attachment of a single β-strand to another single β-strand. Physically, the sum in the exponent of Eq. 5, taken with minus sign, is the ratio between the work W_n to form a fibril of size n and the thermal energy $k_B T$ (k_B is the Boltzmann constant, T is the absolute temperature). For sufficiently dilute solutions, the dimensionless supersaturation s is given by

$$s = \ln \frac{C_1}{C_e}, \quad (7)$$

where C_e (m^{-3}) is the equilibrium concentration of single β-strands in the solution (C_e is also known as the protein solubility). The dimensionless energies ψ and ψ_h per strong (mainly hydrogen) and weak (hydrophobicity-mediated) broken bond are defined by

$$\psi = \frac{E}{2k_B T}, \quad (8)$$

$$\psi_h = \frac{E_h}{2k_B T}. \quad (9)$$

Equations 4 and 5 show that both the equilibrium and the stationary concentrations C_n and X_n of amyloid nanofibrils of a given size n increase strongly with increasing the solution supersaturation s, i.e., with a rise of the actual β-strand concentration C_1 and/or a fall of the protein solubility C_e. As to Eq. 6, in it the proportionality of f_n to l_n reflects the fact that the β-strands are attached predominantly to the ends of the fibril β-sheets (the fibril strong broken bonds are namely at these ends), and the divisor 2 takes into account that a single β-strand has $l_1 = 2$ strong broken bonds.

Combining Eqs. 4–9, we find that when only the nearest-neighbor β-strand interactions are accounted for, the sought stationary fibril size distribution X_n is given by the general ANT formula ($n = 1,2,3,\ldots$) as

$$X_n = C_n \frac{\sum_{k=n}^{M-1}\left(\frac{1}{l_k}\right)\left(\frac{C_e}{C_1}\right)^k \exp(l_k\psi + l_{h,k}\psi_h)}{\sum_{k=1}^{M-1}\left(\frac{1}{l_k}\right)\left(\frac{C_e}{C_1}\right)^k \exp(l_k\psi + l_{h,k}\psi_h)}, \quad (10)$$

in which

$$C_n = C_e\left(\frac{C_1}{C_e}\right)^n \exp[(2 - l_n)\psi + (2 - l_{h,n})\psi_h] \quad (11)$$

is the respective equilibrium fibril size distribution.

Equation 10 is a central result of this study. We emphasize that it is highly reliable, because it is based on the exact

Eq. 4 and because in it the ratio of the two sums is practically unaffected by the approximate character of Eq. 6 for f_n. Importantly, Eq. 11 is in conformity with the law of mass action because of the proportionality of C_n to ${C_1}^n$. Also, as a single β-strand has $l_1 = 2$ strong and $l_{h,1} = 2$ weak broken bonds, Eq. 11 is self-consistent in the sense that at $n = 1$ it returns the identity $C_1 = C_1$.

It should be borne in mind that Eq. 10 is applicable solely to a supersaturated protein solution (i.e., when $C_1 > C_e$), because the stationary size distribution X_n can exist in such a solution only. As then the solution is metastable with respect to fibril nucleation and growth, X_n replaces the equilibrium size distribution C_n that, albeit mathematically well defined, has no physical reality because of the solution metastability. When $C_1 = C_e$ or $C_1 < C_e$, however, the solution is saturated or undersaturated, respectively, and neither nucleation nor growth are possible, for then the solution is thermodynamically stable with respect to these processes. In such a solution, it is the stationary size distribution X_n that is physically irrelevant and the fibril population is described by the equilibrium size distribution C_n from Eq. 5 or Eq. 11.

Changes in fibril solubility due to point mutations

The integrated van' t Hoff equation

$$C_e = C_r \exp\left(-\frac{L}{k_B T}\right), \tag{12}$$

in which C_r (m^{-3}) is a reference protein concentration, relates the equilibrium concentration C_e to the latent heat or enthalpy L (J) (per β-strand) of protein aggregation (e.g., Kashchiev and Auer (51) and Auer and Kashchiev (58)). As found elsewhere (58), in line with the Haas-Drenth lattice model of protein crystals (59), L is approximately given by half of the average binding energy $\langle E_b \rangle$ of a β-strand within the bulk fibrillar phase, i.e., $L \approx \langle E_b \rangle/2$. For the fibril model considered here, we have $\langle E_b \rangle = 2(E + E_h) = 4(\psi + \psi_h)k_B T$, where E, E_h, ψ, and ψ_h are given by Eqs. 1, 2, 8, and 9. Hence, approximately,

$$L = E + E_h = 2(\psi + \psi_h)k_B T. \tag{13}$$

If s_W and s_M are the supersaturations for a wild-type and a point-mutated protein, respectively, employing Eqs. 7 and 12, we find that the change $\Delta s \equiv s_M - s_W$ in the supersaturation upon a point mutation of the wild-type protein at the same temperature T and monomer concentration C_1 is of the form

$$\Delta s = \frac{L_M - L_W}{k_B T}. \tag{14}$$

The latent heat L_W or L_M (per β-strand) of the wild-type protein or the mutant can be calculated from Eq. 13 when the corresponding binding energies E and E_h are known. Also, if $C_{e,W}$ and $C_{e,M}$ are the solubilities of the wild-type and the point-mutated protein, respectively, the mutation-affected change in the protein solubility at the same temperature can be characterized by the ratio

$$\frac{C_{e,M}}{C_{e,W}} = \exp\left(\frac{L_W - L_M}{k_B T}\right). \tag{15}$$

This formula follows from Eq. 12 provided the point mutation has practically no effect on C_r.

RESULTS AND DISCUSSION

Application to Aβ_{40} nanofibrils

Our considerations pertain to that molecular structure proposed by Sachse et al. (60) for Aβ_{40} fibrils, in which the β-strand is formed by a long amino-acid stretch at the C-terminal tail of the protein monomer. As in our previous study (54), we rely on theoretical models (44,45) that identify the most aggregation-prone amino acids in proteins. The theoretical algorithm by Trovato et al. (45) predicts that amino acids 12–20 (VHHQKLVFF) and 31–40 (IIGLMVGGVV) are the most aggregation-prone ones (V = Valine, H = Histidine, Q = Glutamine, K = Lysine, L = Leucine, F = Phenylalanine, I = Isoleucine, G = Glycine, M = Methionine). Joint together, these 19 amino acids form a sequence that defines the β-strand in our nanofibril model (Fig. 1). The sequence-specific energy ε_{jq} of the hydrogen bond between amino acids j and q can vary in the interval $[0, 2\varepsilon]$, depending on the frequency with which two residue types are found paired in neighboring β-strands within a β-sheet in globular proteins (see Eq. 3). For simplicity, we assume that the energy $\varepsilon_{h,jq}$ of a hydrophobicity-mediated bond is the same for all amino-acid pairs, so that $\varepsilon_{h,jq} = \varepsilon_h$, and we set $\varepsilon_h = \varepsilon/10$.

With this set of parameter values, the interstrand binding energies E and E_h for the β-strands (VHHQKLVFFIIGL MVGGVV) arranged parallel within wild-type Aβ_{40} fibrils can be calculated from Eqs. 1–3, and are $E = 25.4\varepsilon$ and $E_h = 1.9\varepsilon$ (see also Table 1). Accordingly, the rounded-up ratio between the β-strand strong and weak binding energies

TABLE 1 Parameter values related to the size distribution of nanosized fibrils of wild-type and point-mutated Aβ_{40} at $T = 300$ K; the corresponding basic hydrogen-bond energy is $\varepsilon = 6.95 \times 10^{-21}$ J

Aβ_{40}	Wild-type	V18N	V18R	V18P
E (J)	1.76×10^{-19}	1.72×10^{-19}	1.68×10^{-19}	1.65×10^{-19}
E_h (J)	1.32×10^{-20}	1.32×10^{-20}	1.32×10^{-20}	1.32×10^{-20}
ψ	21.3	20.8	20.3	19.9
ψ_h	1.59	1.59	1.59	1.59
n_t	14	14	13	13
C_e (μM)	3.9	10.6	28.8	64.1

is $E/E_h = 14$, and the fibril transition size is thus $n_t = 14$ (Table 1) because of our assumption that it is given by the CNT formula (51) $n_t = E/E_h$, which characterizes the fibril equilibrium shape. We note, however, that during nucleation the fibrils are under nonequilibrium conditions: on average, they grow and thus assume a kinetically determined growth shape that may differ considerably from the equilibrium one. Computer simulation data (see Fig. 2 in Zhang and Muthukumar (30)) indicate that the fibril kinetic aspect (length/thickness) ratio that quantifies the fibril growth shape can be several times greater than the thermodynamic aspect ratio E/E_h, i.e., that the fibrils can grow longer than it is thermodynamically expected. It should be kept in mind therefore that, actually, the fibril transition size n_t may have values considerably greater than the value of E/E_h assumed in our analysis.

Aβ_{40} nanofibril size distribution: supersaturation dependence

The nanofibril size distribution for wild-type Aβ_{40} solution at $T = 300$ K and supersaturations $s/\psi_h = 0.9$, 1, 1.2, 2, and 2.5 are displayed in Fig. 3 by lines 0.9, 1, 1.2, 2, and 2.5, respectively. The lines are obtained from Eq. 10 by setting $M = 300$ (as in Cabriolu et al. (52)) and by using the l_n and $l_{h,n}$ values from Table S1 in the Supporting Material and the ψ- and ψ_h-values following from Eqs. 8 and 9 and given in Table 1. According to Eq. 7, with ψ_h from Table 1, the above supersaturations correspond to supersaturation ratios $C_1/C_e = 4.20$, 4.90, 6.78, 24.0, and 53.3, respectively.

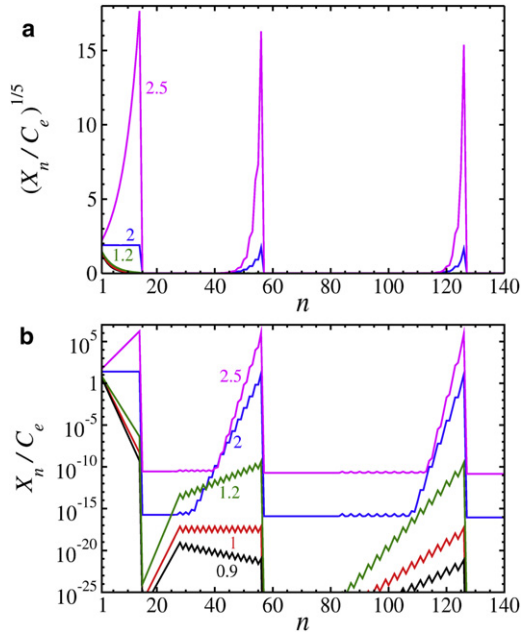

FIGURE 3 Stationary nanofibril size distribution X_n on (*a*) $X_n^{1/5}$ and (*b*) log X_n scale according to Eq. 10 at $T = 300$ K and supersaturations $s/\psi_h = 0.9$, 1, 1.2, 2, and 2.5 (as indicated) for wild-type Aβ_{40}.

As seen in Fig. 3, Eq. 10 predicts a series of peaks, each of them being at a certain fixed fibril size for high enough supersaturations and at a smaller, but again fixed, size for sufficiently low supersaturations. The first peak is positioned either at fibril size n'_1 or at fibril size $n''_1 > n'_1$ that, due to our choice of the nanofibril evolution mechanism illustrated in Fig. 2, are given by $n'_1 = 1$ (for $0 < s < 2\psi_h$) and by $n''_1 = n_t$ (for $s > 2\psi_h$), where $n_t = 14$ and $2\psi_h = 3.18$ because of the ψ- and ψ_h-values used in our calculation of X_n in Fig. 3. At $s = 2\psi_h$ (line 2), the peak is plateaulike, extending from $n = n'_1$ to $n = n''_1$, i.e., from the monomer size to the fibril size n_t for transition of the fibril from one-dimensional to two-dimensional aggregate.

Similarly, the fibril sizes n'_j and n''_j at which the second, third, etc. peaks are positioned at low or high supersaturations, respectively, are obtainable from the relations ($j = 2,3,4,\ldots$),

$$n'_j = j(j-1)n_t \tag{16}$$

(for $0 < s < 2\psi_h/j$) and

$$n''_j = j^2 n_t \tag{17}$$

(for $s > 2\psi_h/j$). At $s = 2\psi_h/j$, the j^{th} peak is a plateau (jagged for $j = 2,3,4,\ldots$) that extends from $n = n'_j$ to $n = n''_j$. In Fig. 2, the first three nanofibrils of size n'_j and of size n''_j are colored pink and blue, respectively.

Fig. 2 and Eqs. 16 and 17 show that both the n'_j-sized and the n''_j-sized fibrils are constituted of j β-sheets, but their lengths m'_j and m''_j, respectively, are multiples of n_t (except $m'_1 = 1$):

$$m'_j = (j-1)n_t, \tag{18}$$

$$m''_j = jn_t. \tag{19}$$

Equations 16 and 17 show also that for $s > 2\psi_h$, i.e., for high enough supersaturations, all peaks (including the first one) remain fixed at characteristic or "magic" fibril sizes n''_j given by Eq. 17. The n''_j-sized fibril itself is j β-sheets thick and jn_t β-strands long (see the *blue fibrils* in Fig. 2), i.e., its length m''_j is determined by Eq. 19. This equation implies that when $s > 2\psi_h$, if plotted as a function of the nanofibril length m, the stationary size distribution of the nanofibrils will be multipeaked at magic lengths m''_j, the peaks being at the same distance n_t from their nearest neighbors on the m axis. We note, however, that this result about equidistanced peaks in the nanofibril length distribution characterizes the nanofibril shape sequence illustrated in Fig. 2. Deviations from this sequence, in particular regarding the length at which the nanofibril acquires the energetically very costly β-strand that gives birth of its subsequent

β-sheet, could lead to different interpeak distances on the nanofibril length axis.

Combining Eq. 17 with the formula $n^*_j = j^2 n_t + 1$ for the successive nucleus sizes n^*_j (52), we arrive at the equation ($j = 1,2,3,4,\ldots$)

$$n_j^* = n_j'' + 1, \tag{20}$$

which says that the j^{th} fibril nucleus is just one β-strand bigger than the fibril corresponding to the j^{th} peak at high enough supersaturations (see Fig. 2 in which the first three fibril nuclei are hatched, and the n''_j-sized fibrils are *blue*). Equation 20 is of particular interest, for it seems to be a result of general validity when amyloid fibrils are nucleated by the mechanism of direct polymerization of β-strands and when $s > 2\psi_h$. According to this equation, with a negligible difference of one β-strand, the positions n''_j of the peaks in a nanofibril size distribution that is obtained in a real or computer experiment at a high enough supersaturation s merely represent the sizes n^*_j of the successive fibril nuclei.

Table S2 lists the first three values of n'_j, m'_j, n''_j, and m''_j in the respective s ranges. To get a feel of these values for the Aβ_{40} protein, we recall that the molecular mass of the protein monomer is 4.3 kDa (61), and that the distance between β-strands in the Aβ_{40} fibril is ~0.47 nm (62). Hence, from Eqs. 17 and 19 with $n_t = 14$, the first three peaks of the nanofibril mass or length distribution are at magic masses 4.3 n''_j of 60.2, 241, and 542 kDa or lengths 0.47 m''_j of 6.58, 13.2, and 19.7 nm provided the solution is at supersaturation $s > 2\psi_h = 3.18$.

The physical reason for which the size distribution X_n of nanosized fibrils is peaked at magic sizes n'_j or n''_j is the presence of deep local minima of the work W_n to form a fibril of n β-strands (see Fig. 3*b* of Cabriolu et al. (52)). The attachment of a β-strand to the end of an n'_j-sized fibril or to the surface of an n''_j-sized fibril requires considerable work so that the n'_j- or n''_j-sized fibrils become the most numerous in the solution and thus give rise to peaks in the nanofibril size distribution. As to the s value at which the j^{th} peak is plateaulike, it coincides with that at which, according to CNT (Kashchiev and Auer (51), Eq. 13), a $j\beta$-sheet is in thermodynamic equilibrium with the solution, i.e., it neither grows nor dissolves. We note as well that, regardless of the s value, a given peak concentration $X_{p,j}$ is always higher than the subsequent concentration $X_{p,j+1}$, i.e., the peak concentrations of the nanofibrils obey the order $X_{p,1} \geq X_{p,2} \geq X_{p,3} \geq \ldots$.

Clearly, X_n will be closer to a monotonically decreasing function of n when the work $(-s + 2\psi)k_BT$ (Kashchiev and Auer (51)) of thicknesswise β-strand attachment to the kinkless surface of any of the n''_j-sized fibrils (Fig. 2) is sufficiently smaller than the thermal energy k_BT. Thus, the condition for relatively low peaks in the stationary nanofibril size distribution reads

$$2\psi < 1 + s. \tag{21}$$

This condition shows that, as usually $\psi > 1$, well-pronounced peaks at magic sizes (or lengths) are to be expected in the stationary size (or length) distributions of most amyloid nanofibrils (the effect of the supersaturation is quite small, because amyloid fibril nucleation occurs at s values considerably smaller than those of ψ (51,52)). For nanofibrils of shorter β-strands the peaks will be less pronounced, because ψ diminishes with decreasing β-strand length. In principle, any biophysicochemical factor that lessens the ψ-value will also contribute to the suppression of the X_n peaks and, hence, to the possible disappearance of the magic nanofibril sizes or lengths.

Preferred sizes, lengths, or thicknesses of Aβ oligomers or fibrils have been recorded in many experiments (e.g., (29,61–65)). In this respect, our finding of magic sizes is in qualitative agreement with experiment. No experimental data we are aware of, however, allows us using it for a quantitative verification of this finding. Reported histograms of the fibril mass-per-length ratio (60,61,63–69) exhibit one or more peaks, but pertain to portions of rather long fibrils and show that these usually contain an integer number of β-sheets. To reliably quantify the magic nanofibril sizes, we need histograms of the masses of entire fibrils with nanoscale lengths. To the best of our knowledge, only Goldsbury et al. (61) have obtained such a histogram (Fig. 4*D* in their article), but it is for pseudospherical rather than fibrillar Aβ_{40} aggregates and the authors could fit it by two Gaussian curves with maxima at masses of 311 and 554 kDa corresponding to 72 and 128 protein molecules. The experimental size distributions of Garai et al. (29) for solutions of Aβ_{40} also exhibit peaks at magic sizes represented by the aggregate hydrodynamic radii. Again, a quantitative comparison between our theory and the experiment of Garai et al. (29) is not possible. The reported size distributions are not stationary, the aggregate morphology is unknown, and the second peak of the size distributions is at radius of ~50 nm. This radius corresponds to the rather large number of 1.2×10^5 protein molecules if, as in Garai et al. (29), one uses 4.3 nm^3 for the Aβ_{40} molecular volume and assumes that the aggregates are spherical.

Very interesting are as well the experimental data of Kellermayer et al. (62), which support our finding of magic nanofibril lengths. These authors have presented a histogram of most frequently observed increments of the lengths of epitaxially growing individual A$\beta_{25\text{-}35}$ fibrils (the *inset* in Fig. 2*a* of Kellermayer et al. (62)). The histogram exhibits five successive peaks at 6.5, 13.3, 23.2, 32.5, and 40 nm. As noted by Kellermayer et al. (62), the nearly 7-nm interpeak distance "corresponds to a stretch of fibril containing $\approx$15 peptides along its length." This is in close quantitative agreement with our finding for peaks in the Aβ_{40} nanofibril length distribution, which are equidistanced at 14 β-strands or, equivalently, 6.6 nm. However,

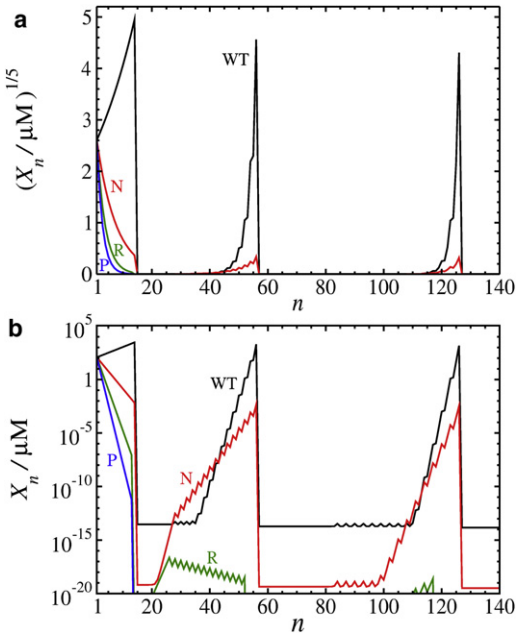

FIGURE 4 Stationary nanofibril size distribution on (*a*) $X_n^{1/5}$ and (*b*) log X_n scale according to Eq. 10 at $T = 300$ K and monomer concentration $C_1 = 120$ μM for wild-type Aβ_{40} and its mutants V18N, V18R, and V18P: (*black*) line WT, wild-type; (*red*) line *N*, V18N; (*green*) line *R*, V18R; and (*blue*) line *P*, V18P.

this agreement is very likely to be fortuitous, because whereas the fibrils of Kellermayer et al. (62) elongate at a practically constant thickness, according to our model (Fig. 2) every subsequent fibril lengthening of n_t β-strands is preceded by fibril thickening of one β-sheet.

Aβ_{40} nanofibril size distribution: effect of mutations

The presented theoretical framework enables us also to study the effect of point mutations on the nanofibril size distribution. In particular, at fixed T and C_1, we examine how X_n is affected by the point mutations of residue 18 (V) with amino acids Asparagine (N), Arginine (R), and Proline (P), because the effect of these mutations of Aβ_{40} on the fibrillation kinetics has been investigated experimentally (35).

To be able to calculate X_n from Eq. 10 for each mutant, we need to know the corresponding fibril solubility C_e and energy parameters ψ and ψ_h that have different values for each mutant, because the binding energies E and E_h of the β-strands within the fibrils are different (see Table 1). For the wild-type Aβ_{40}, we employ the experimentally obtained $C_e = 3.9$ μM at $T = 300$ K (29). As C_e has not been determined experimentally for any of the mutants, we find it theoretically by making use of Eqs. 13 and 15 and the ψ- and ψ_h-values from Table 1. The C_e values computed in this way for fibrils with point mutations V18N, V18R, and V18P are given in Table 1. Using the procedure outlined above for the wild-type Aβ_{40}, from Eq. 10 we calculate X_n for each mutant by setting $M = 300$ (as in Cabriolu et al. (52)) and by using the ψ- and ψ_h-values from Table 1 and the l_n and $l_{h,n}$ values from Table S1 and Table S3. For mutations V18N, V18R, and V18P the transition size $n_t = E/E_h$ has the values 14, 13, and 13, respectively, and is equal or nearly equal to the transition size $n_t = 14$ for the wild-type protein (Table 1).

The nanofibril size distributions obtained at $T = 300$ K and $C_1 = 120$ μM are shown in Fig. 4. At this concentration, the supersaturation for the wild-type protein and the three mutants V18N, V18R, and V18P can be calculated from Eq. 7 and is $s = 3.43$, 2.43, 1.43, and 0.63, respectively. As can be seen in Fig. 4, the main effect of the mutations is that the peaks for all magic nanofibril sizes n'_j or n''_j lose height. This is so because by considerably increasing the fibril solubility, the mutations lower the solution supersaturation, which leads to a strong decrease in the entire fibril population and, accordingly, in the peak heights. In contrast, the effect of the mutations on n'_j or n''_j is minor (because n_t changes only from 14 to 13 for two of the mutants) and, as a result, the peak positions remain essentially the same. For mutant V18R, the second peak is at $n'_2 = 26$, because the C_e value for this mutant is so great that the corresponding supersaturation ratio C_1/C_e is too low for the peak to be at $n''_2 = 52$.

The effect of mutations on the fibril population could be exemplified by considering the number of nanofibrils of 13 β-strands in a solution of volume $V = 100$ mL typically used in experiments. From Eq. 10 (see also Fig. 4) we calculate that whereas the number $X_{13}V$ of wild-type fibrils of size $n = 13$ is 10^{20}, the V18N and V18R fibrils of the same size are much less numerous: the V18N ones are 9×10^{14}, and the V18R ones are 10^{10}. Remarkably, the fibrils of the same size for mutant V18P are practically absent from the solution, because their number is 4×10^5. It is worth noting also that the decrease of the peak heights for the different mutants that is seen in Fig. 4 is consistent with experiments on protein fibrillation kinetics (35), which reveal that the lag time before detectable fibrillation is longer for the mutants V18N and V18R than for the wild-type Aβ_{40}; and that the mutant V19P is not detectable at all. This consistency is in support of the correlation reported earlier (54) between ANT nucleation rates of amyloid fibrils and experimentally measured fibrillation lag-times of wild-type and point-mutated Aβ proteins.

CONCLUSION

The analysis made illustrates the ability of the nucleation theory to describe the size distribution of amyloid nanofibrils in protein solutions when nucleation occurs by the mechanism of direct polymerization of β-strands into β-sheets. The atomistic modeling of the nanofibrils of successively layered β-sheets leads to a general formula

(Eq. 10) for the stationary size distribution as a function of the protein concentration and of the temperature (via C_e, ψ, and ψ_h that, given the nanofibril shape sequence, are the only three theoretical parameters). Application of this formula to Aβ_{40} nanofibrils reveals the existence of a series of characteristic peaks in the size distribution, which are positioned at magic nanofibril sizes (or lengths) and which are due to the sufficiently deep local minima in the work for fibril formation. This finding of magic sizes or lengths is consistent with experimental results for the size distribution of aggregates in solutions of Aβ_{40} proteins.

We believe our analysis provides a new opportunity to obtain information about the sizes of the successive fibril nuclei, because according to Eq. 20, these sizes are practically equal to the magic sizes at which the nanofibril size distribution is peaked when the solution supersaturation is sufficiently high. Also, we believe our atomistic approach makes it possible to gain novel insight into the effect of point mutations on the size distribution of amyloid nanofibrils, an effect that may play a role in experimentally observed substantial differences in the fibrillation lag-times of wild-type and point-mutated Aβ proteins.

The results obtained remain without quantitative verification because of the lack of suitable experimental or simulation data for the size distribution of amyloid nanofibrils. It is hoped that our study will inspire new real and computer experiments especially aimed at investigating this distribution and/or the magic nanofibril sizes or lengths.

SUPPORTING MATERIAL

Three tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(11)01184-2.

This work was supported by Engineering and Physical Science Research Council grant No. EP/G026165/1.

REFERENCES

1. Chiti, F., and C. M. Dobson. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–366.
2. Gazit, E. 2007. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* 36:1263–1269.
3. Sunde, M., and C. Blake. 1997. The structure of amyloid fibrils by electron microscopy and x-ray diffraction. *Adv. Protein Chem.* 50:123–159.
4. Jaroniec, C. P., C. E. MacPhee, …, R. G. Griffin. 2004. High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. USA.* 101:711–716.
5. Makin, O. S., E. Atkins, …, L. C. Serpell. 2005. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. USA.* 102:315–320.
6. Tycko, R. 2006. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys.* 39:1–55.
7. Sawaya, M. R., S. Sambashivan, …, D. Eisenberg. 2007. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature.* 447:453–457.
8. Wasmer, C., A. Lange, …, B. H. Meier. 2008. Amyloid fibrils of the HET-s(218-289) prion form a β-solenoid with a triangular hydrophobic core. *Science.* 319:1523–1526.
9. Serio, T. R., A. G. Cashikar, …, S. L. Lindquist. 2000. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science.* 289:1317–1321.
10. Auer, S., C. M. Dobson, and M. Vendruscolo. 2007. Characterization of the nucleation barriers for protein aggregation and amyloid formation. *HFSP J.* 1:137–146.
11. Auer, S., F. Meersman, …, M. Vendruscolo. 2008. A generic mechanism of emergence of amyloid protofilaments from disordered oligomeric aggregates. *PLOS Comput. Biol.* 4:e1000222.
12. Hofrichter, J., P. D. Ross, and W. A. Eaton. 1974. Kinetics and mechanism of deoxyhemoglobin S gelation: a new approach to understanding sickle cell disease. *Proc. Natl. Acad. Sci. USA.* 71:4864–4868.
13. Ferrone, F. A., J. Hofrichter, and W. A. Eaton. 1985. Kinetics of sickle hemoglobin polymerization. I. Studies using temperature-jump and laser photolysis techniques. *J. Mol. Biol.* 183:591–610.
14. Ferrone, F. A., J. Hofrichter, and W. A. Eaton. 1985. Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol.* 183:611–631.
15. Hofrichter, J. 1986. Kinetics of sickle hemoglobin polymerization. III. Nucleation rates determined from stochastic fluctuations in polymerization progress curves. *J. Mol. Biol.* 189:553–571.
16. Jarrett, J. T., and P. T. Lansbury, Jr. 1993. Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell.* 73:1055–1058.
17. Lomakin, A., D. S. Chung, …, D. B. Teplow. 1996. On the nucleation and growth of amyloid β-protein fibrils: detection of nuclei and quantitation of rate constants. *Proc. Natl. Acad. Sci. USA.* 93:1125–1129.
18. Lomakin, A., D. B. Teplow, …, G. B. Benedek. 1997. Kinetic theory of fibrillogenesis of amyloid β-protein. *Proc. Natl. Acad. Sci. USA.* 94:7942–7947.
19. Harper, J. D., and P. T. Lansbury, Jr. 1997. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* 66:385–407.
20. Galkin, O., and P. G. Vekilov. 2004. Mechanisms of homogeneous nucleation of polymers of sickle cell anemia hemoglobin in deoxy state. *J. Mol. Biol.* 336:43–59.
21. Nguyen, H. D., and C. K. Hall. 2004. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl. Acad. Sci. USA.* 101:16180–16185.
22. Nguyen, H. D., and C. K. Hall. 2005. Kinetics of fibril formation by polyalanine peptides. *J. Biol. Chem.* 280:9074–9082.
23. Nguyen, H. D., and C. K. Hall. 2006. Spontaneous fibril formation by polyalanines; discontinuous molecular dynamics simulations. *J. Am. Chem. Soc.* 128:1890–1901.
24. Ferrone, F. A. 2006. Nucleation: the connections between equilibrium and kinetic behavior. *Methods Enzymol.* 412:285–299.
25. Galkin, O., R. L. Nagel, and P. G. Vekilov. 2007. The kinetics of nucleation and growth of sickle cell hemoglobin fibers. *J. Mol. Biol.* 365:425–439.
26. Linse, S., C. Cabaleiro-Lago, …, K. A. Dawson. 2007. Nucleation of protein fibrillation by nanoparticles. *Proc. Natl. Acad. Sci. USA.* 104:8691–8696.
27. Auer, S., C. M. Dobson, …, A. Maritan. 2008. Self-templated nucleation in peptide and protein aggregation. *Phys. Rev. Lett.* 101:258101.
28. Xue, W. F., S. W. Homans, and S. E. Radford. 2008. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proc. Natl. Acad. Sci. USA.* 105:8926–8931.
29. Garai, K., B. Sahoo, …, S. Maiti. 2008. Quasihomogeneous nucleation of amyloid β yields numerical bounds for the critical radius, the surface tension, and the free energy barrier for nucleus formation. *J. Chem. Phys.* 128:045102.

30. Zhang, J., and M. Muthukumar. 2009. Simulations of nucleation and elongation of amyloid fibrils. *J. Chem. Phys.* 130:035102.

31. Nielsen, L., R. Khurana, …, A. L. Fink. 2001. Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry.* 40:6036–6046.

32. Kim, Y. S., S. P. Cape, …, J. F. Carpenter. 2001. Counteracting effects of renal solutes on amyloid fibril formation by immunoglobulin light chains. *J. Biol. Chem.* 276:1626–1633.

33. Zhu, L., X. J. Zhang, …, S. Perrett. 2003. Relationship between stability of folding intermediates and amyloid formation for the yeast prion Ure2p: a quantitative analysis of the effects of pH and buffer system. *J. Mol. Biol.* 328:235–254.

34. Kunes, K. C., D. L. Cox, and R. R. P. Singh. 2005. One-dimensional model of yeast prion aggregation. *Phys. Rev. E.* 72:051915.

35. Christopeit, T., P. Hortschansky, …, M. Fändrich. 2005. Mutagenic analysis of the nucleation propensity of oxidized Alzheimer's β-amyloid peptide. *Protein Sci.* 14:2125–2131.

36. Hortschansky, P., V. Schroeckh, …, M. Fändrich. 2005. The aggregation kinetics of Alzheimer's β-amyloid peptide is controlled by stochastic nucleation. *Protein Sci.* 14:1753–1759.

37. Pedersen, J. S., J. M. Flink, …, D. E. Otzen. 2006. Sulfates dramatically stabilize a salt-dependent type of glucagon fibrils. *Biophys. J.* 90:4181–4194.

38. Grudzielanek, S., V. Smirnovas, and R. Winter. 2006. Solvation-assisted pressure tuning of insulin fibrillation: from novel aggregation pathways to biotechnological applications. *J. Mol. Biol.* 356:497–509.

39. Fändrich, M. 2007. Absolute correlation between lag time and growth rate in the spontaneous formation of several amyloid-like aggregates and fibrils. *J. Mol. Biol.* 365:1266–1270.

40. Routledge, K. E., G. G. Tartaglia, …, S. E. Radford. 2009. Competition between intramolecular and intermolecular interactions in an amyloid-forming protein. *J. Mol. Biol.* 389:776–786.

41. Hellstrand, E., B. Boland, …, S. Linse. 2009. Amyloid Î²-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *Am. Chem. Soc. Chem. Neurosci.* 1:13–18.

42. Chiti, F., M. Stefani, …, C. M. Dobson. 2003. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature.* 424:805–808.

43. Yoon, S., and W. J. Welsh. 2004. Detecting hidden sequence propensity for amyloid fibril formation. *Protein Sci.* 13:2149–2160.

44. Fernandez-Escamilla, A. M., F. Rousseau, …, L. Serrano. 2004. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22:1302–1306.

45. Trovato, A., F. Chiti, …, F. Seno. 2006. Insight into the structure of amyloid fibrils from the analysis of globular proteins. *PLOS Comput. Biol.* 2:e170.

46. Tartaglia, G. G., A. P. Pawar, …, M. Vendruscolo. 2008. Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* 380:425–436.

47. Vestergaard, B., M. Groenning, …, D. I. Svergun. 2007. A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *PLoS Biol.* 5:e134.

48. Oliveira, C. L., M. A. Behrens, …, J. S. Pedersen. 2009. A SAXS study of glucagon fibrillation. *J. Mol. Biol.* 387:147–161.

49. Garai, K., R. Sureka, and S. Maiti. 2007. Detecting amyloid-β aggregation with fiber-based fluorescence correlation spectroscopy. *Biophys. J.* 92:L55–L57.

50. Auer, S., and D. Kashchiev. 2010. Insight into the correlation between lag time and aggregation rate in the kinetics of protein aggregation. *Proteins Struct. Funct. Bioinform.* 78:2412–2416.

51. Kashchiev, D., and S. Auer. 2010. Nucleation of amyloid fibrils. *J. Chem. Phys.* 132:215101.

52. Cabriolu, R., D. Kashchiev, and S. Auer. 2010. Atomistic theory of amyloid fibril nucleation. *J. Chem. Phys.* 133:225101.

53. Schmit, J. D., K. Ghosh, and K. Dill. 2011. What drives amyloid molecules to assemble into oligomers and fibrils? *Biophys. J.* 100:450–458.

54. Cabriolu, R., and S. Auer. 2011. Amyloid fibrillation kinetics: insight from atomistic nucleation theory. *J. Mol. Biol.* 411:275–285.

55. Kashchiev, D. 2008. Toward a better description of the nucleation rate of crystals and crystalline monolayers. *J. Chem. Phys.* 129:164701.

56. Fersht, A. R., J.-P. Shi, …, G. Winter. 1985. Hydrogen bonding and biological specificity analyzed by protein engineering. *Nature.* 314:235–238.

57. Kashchiev, D. 2000. Nucleation: Basic Theory with Applications. Butterworth Heinemann, Oxford, UK.

58. Auer, S., and D. Kashchiev. 2010. Phase diagram of α-helical and β-sheet forming peptides. *Phys. Rev. Lett.* 104:168105.

59. Haas, C., and J. Drenth. 1995. The interaction energy between two protein molecules related to physical properties of their solution and their crystals and implications for crystal growth. *J. Cryst. Growth.* 154:126–135.

60. Sachse, C., M. Fändrich, and N. Grigorieff. 2008. Paired β-sheet structure of an Aβ (1-40) amyloid fibril revealed by electron microscopy. *Proc. Natl. Acad. Sci. USA.* 105:7462–7466.

61. Goldsbury, C. S., S. Wirtz, …, P. Frey. 2000. Studies on the in vitro assembly of a β 1-40: implications for the search for a β-fibril formation inhibitors. *J. Struct. Biol.* 130:217–231.

62. Kellermayer, M. S., A. Karsai, …, B. Penke. 2008. Stepwise dynamics of epitaxially growing single amyloid fibrils. *Proc. Natl. Acad. Sci. USA.* 105:141–144.

63. Antzutkin, O. N., R. D. Leapman, …, R. Tycko. 2002. Supramolecular structural constraints on Alzheimer's β-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. *Biochemistry.* 41:15436–15450.

64. Goldsbury, C., P. Frey, …, S. A. Müller. 2005. Multiple assembly pathways underlie amyloid-β fibril polymorphisms. *J. Mol. Biol.* 352:282–298.

65. Petkova, A. T., Y. Ishii, …, R. Tycko. 2002. A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. USA.* 99:16742–16747.

66. Schmidt, M., C. Sachse, …, N. Grigorieff. 2009. Comparison of Alzheimer Aβ (1-40) and Aβ (1-42) amyloid fibrils reveals similar protofilament structures. *Proc. Natl. Acad. Sci. USA.* 106:19813–19818.

67. Chen, B., K. R. Thurber, …, R. Tycko. 2009. Measurement of amyloid fibril mass-per-length by tilted-beam transmission electron microscopy. *Proc. Natl. Acad. Sci. USA.* 106:14339–14344.

68. Komatsu, H., E. Feingold-Link, …, P. H. Axelsen. 2010. Intrinsic linear heterogeneity of amyloid β protein fibrils revealed by higher resolution mass-per-length determinations. *J. Biol. Chem.* 285:41843–41851.

69. Goldsbury, C., U. Baxa, …, S. A. Müller. 2011. Amyloid structure and assembly: insights from scanning transmission electron microscopy. *J. Struct. Biol.* 173:1–13.